

RESEARCH

Open Access



Game theoretical analysis of rate adaptation protocols conciliating QoS and QoE

Smrati Gupta^{1,4*}, E. V. Belmega^{2,3} and M. A. Vázquez-Castro¹

Abstract

The recent increase in the use of wireless networks for video transmission has led to the increase in the use of rate-adaptive protocols to maximize the resource utilization and increase the efficiency in the transmission. However, a number of these protocols lead to interactions among the users that are subjective in nature and affect the overall performance. In this paper, we present an in-depth analysis of interplay between the wireless network dynamics and video transmission dynamics in the light of subjective perceptions of the end users in their interactions. We investigate video exchange applications in which two users interact repeatedly over a wireless relay channel. Each user is driven by three conflicting objectives: maximizing the Quality of Service (QoS) and Quality of Experience (QoE) of the received video, while minimizing the transmission cost. Non-cooperative repeated games model precisely interactions among users with independent agendas. We show that adaptive video exchange is impossible if the duration of the interaction is determined. However, if the users interact indefinitely, they achieve cooperation via exchange of video streams. Our simulations evidence the tradeoff between users' QoS and QoE of their received video. The expected duration of the interaction plays a role and draws the region of solution trade-offs. We propose further means of shaping this region using Pareto optimality and user-fairness arguments. This work proposes a concrete game theoretical framework that allows the optimal use of traditional protocols by taking into account the subjective interactions that occur in practical scenarios.

Keywords: Rate adaptation, QoS vs. QoE tradeoff, Video exchange, Non-cooperative repeated games

1 Introduction

The past decade has seen an enormous growth in users of wireless networks. With the increasing variety of video applications via wireless channels, it is imperative that this number will grow. Given the limited resources at their disposal, competition emerges among the users to access these video services effectively. This competition affects the experienced video quality and plays a role in optimizing the resource allocation.

An important demand of the users of the video applications is high quality of experience of the perceived video. The QUALITY of Experience (QoE) at the end user depends on the temporal and spatial structure of

the video, and its optimization involves the minimization of video distortion and of disruptions in video playback. The temporal structure may relate to many features, for example, to the continuity in the video observed at the end user without freezing or disruptions, ideally implying timely playback of the video. Similarly, the spatial structure may relate to features, for example, to the avoidance of formation of artifacts due to packet losses at the end user [1]. Furthermore, the wireless nodes involved in the video transmission are distributed and autonomous devices. In a video exchange scenario between two selfish autonomous users, the ability of a user to obtain a desired quality video depends on the action chosen by the other user, i.e. how much information it sends. In turn, the other user will incur a transmission cost. This leads to a natural conflict between minimizing their own transmission cost while maximizing the Quality of Service (QoS) and QoE of their video. This paper addresses such competitive video exchange between two users.

*Correspondence: smrati.gupta@ca.com

Preliminary results of this work have been presented at the ASMS/SPSC conference in 2014.

¹Department of Telecommunications and Systems Engineering, Autonomous University of Barcelona, Cerdanyola del Vallès, Spain

⁴CA Strategic Research, CA Technologies, Barcelona, Spain

Full list of author information is available at the end of the article

A relevant framework to model this type of competitive interactions is provided by game theory. Tools from game theory have already been used to study various aspects of transmission over wireless networks [2, 3]. To be more precise, the game theoretical framework has been applied at the physical layer to design power allocation policies [4, 5] and at the application layer to design rate allocation policies [6, 7]. Furthermore, other types of strategic interactions such as network topology selection games [8], pricing games between service providers and users for network congestion control [9]. The aforementioned works consider a one-shot interaction model; however, repeated interactions that take place over multiple stages seems more realistic. Repeated games have been used in [10] to study a distributed state estimation problem for the inter-connected electrical power grid; also, in [11], the authors study a distributed power control problem in a wireless communication scenario. Therefore, the repeated games' framework seems to be very promising for many other interactive multi-user applications such as distributed rate adaptation problems.

The game theoretical analysis of different transmission protocols has been considered in the literature [7, 12]. In [7], the authors have modeled the players as the Transmission Control Protocol (TCP) flows and the actions as the rate adaptation parameters to optimize their own average throughputs. The user actions, which are the rate adaptation parameters, control the number of buffer packets that are stored to be transmitted at a later stage. In [12], the competition in TCP has been modeled by allowing the users to choose the gradient of rate adaptation. In these works, the players have discrete (often binary) sets of actions. It is important to build models that allow the users to choose smoother actions, e.g. from continuous sets.

The main contributions of this work are as follows:

- We present a video exchange interaction among two users by a non-cooperative repeated game framework. The overall game-theoretical framework that has been built is particularly suited for wireless networking analysis. The reason is that both the wireless channel dynamics (network resource) and the video rate dynamics can be jointly treated analytically to analyze the performance of interactive protocols. Moreover, the type of analysis based on selection of (conflicting) objectives allows to account for a wide range of pure technological issues to be studied against different subjective considerations.
- We show the application of the game theoretical framework to the analysis of the concrete problem of QoE-driven rate adaptation over underlying Markovian channel dynamics taking into account

subjective factors. The use of this framework has revealed the existence of a trade-off between QoE and QoS in such systems. This trade-off arises due to conflicting paradigms that are desired by the users: increase in the quality of the video, or QoS, received leading to higher rate of adaptation to channel conditions and increase in quality of video playback, or QoE that is inversely affected by faster rate adaptation due to higher probabilities of exceeding channel capacity. This trade-off captures the three-fold impact of dynamics among the network resources, video transmission and subjective interactions, all three of which are modeled in our framework.

- We illustrate that this trade-off has been found to be controlled by not only the expected channel conditions but other factors like the expected duration of the interaction among the users and the tolerance of the users and Pareto optimality arguments. As the expected duration of interaction increases, the speed of rate adaptation to the channel conditions increases providing better QoS but leading to increase in outages inversely affecting the QoE. Likewise, it is also shown and quantized that the achievable QoS and QoE of the user depends on the level of tolerance towards the performance gains of the other user exceeding its own. Further, we argue that rational users are more likely to agree upon Pareto optimal trade-offs. Pareto optimality of an outcome means that no other tradeoff exists that offers a strictly better utility to one user without decreasing the other's. Hence, a number of factors are shown to directly or indirectly determine the achievable rates and consequently the QoS and QoE obtained by the users.
- Overall, our results indicate that competing protocols over underlying wireless networks can be more accurately understood by considering also subjective metrics. Robust protocols for wireless channels should then be able to take into account underlying trade-offs for optimal and stable behavior, etc.

This work extends and improves the conference version [13] in four aspects. (i) The results regarding the solution of the repeated game were only announced in [13]. Here, we provide the complete proofs of all our claims. (ii) We extend the system model to include both satellite- and terrestrial-based bent-pipe topology. (iii) We add a discussion on the means to shape the optimal trade-off region of the game using user-fairness and Pareto optimality arguments. (iv) At last, our simulation section is wider and offers a more complete view on how the system parameters impact the outcome of the game with an emphasis on the QoS vs. QoE trade-offs.

This paper is organized as follows. In Section 2, we describe the system model and the QoE-driven rate adaptation performance metrics which are used in the paper. We present the game theoretical framework for video transmission in Section 3. The numerical results are presented in Section 4 followed by the conclusions in Section 5.

2 System and rate adaptation models

In this section, we describe the system model and the rate adaptation scheme. We also define the users performance metrics.

2.1 System model

2.1.1 Topology

The system under study is composed of two nodes or users that may exchange their videos via a relay node. This model relates to video streaming situations between two parties across geographically separated areas. User A (or B) transmits the video packets to the relay, and the relay forwards these packets the other user B (or A) as shown in Fig. 1.

2.1.2 Channel model

The wireless channel between any two nodes has variable capacity to transmit packets depending on factors like environmental conditions (rain, cloud, etc.), and congestion due to other users. We model channel links between the nodes to have a certain throughput capacity which varies with time. The throughput capacity is the maximum number of packets that can be transmitted between any node k and node l at time t and is denoted by $R_{kl}(t)$. It is measured in packets per time slot (pps). As shown in Fig. 1, $R_{AB} = 0$ for all t . All links are assumed to have equal throughput capacities $R_c(t)$, that is, $R_{kl}(t) = R_c(t)$, $\forall kl \in \{AR, BR, RA, RB\}$ at any time instant t . Each time instant is described by $t = n\Delta$, where $n \in \mathbb{Z}^+$ and Δ is the interval size, i.e. a precision parameter. We will

refer to instantaneous quantities using the integer value n henceforth. We assume that the relay node (e.g. satellite) simply forwards the packet without any processing. Therefore, $R_c(n)$ is the instantaneous end-to-end throughput capacity of the channel between A and B and vice versa.

This assumption (of the symmetric channel links) is taken for the sake of simplicity of analysis and illustration. The main results of this work can be generalized to the case of asymmetric links. However, the underlying mathematical analysis becomes more tedious and hinges the reader into the understanding of the major ideas and intuition behind the repeated user interactions.

We consider a bent-pipe topology that belongs to a bigger and denser network. Our aim is to study the dynamics of transmission within this subset of a denser network such that the results can be modularized and extrapolated to more general cases.

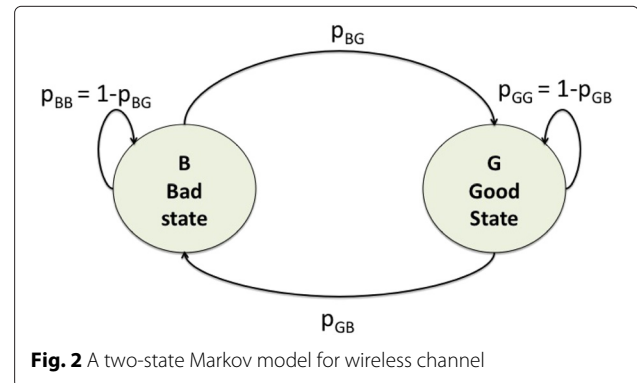
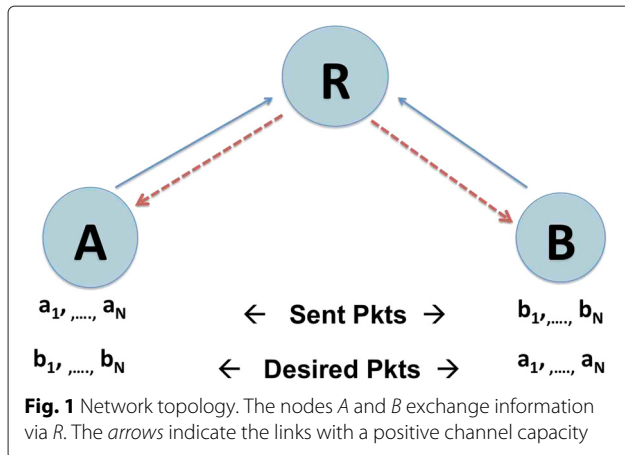
The underlying wireless channel is modeled by a two-state Markov random process as described in [14] and shown in Fig. 2. The “good” state corresponds to the channel being in favorable conditions (e.g., line of sight and clear weather). The “bad” state refers to the channel being in unfavorable conditions (e.g. obstructive conditions like clouds and raining). The probability of the channel to move from a good to a bad state and vice versa is p_{GB} and p_{BG} , respectively. Consequently, the probability to remain in good and bad state are $p_{GG} = 1 - p_{GB}$ and $p_{BB} = 1 - p_{BG}$. From this model, we can infer the probability of the channel to be in bad and good states (i.e., p_B and p_G):

$$p_B = \frac{p_{GB}}{p_{GB} + p_{BG}} \quad (1)$$

$$p_G = \frac{p_{BG}}{p_{GB} + p_{BG}}. \quad (2)$$

When in a good state, the channel has a throughput capacity of R_1 , whereas, in a bad state, the capacity is lower $R_2 < R_1$. Therefore, at any time instant $t = n\Delta$,

$$R_c(n) = \begin{cases} R_G & \text{good state} \\ R_B & \text{bad state} \end{cases} \quad (3)$$



The channel coherence time is given by $T_c = c\Delta$, $c \in \mathbb{Z}^+$. We assume that $c \gg 1$ which means that the channel coherence time is very large which is reasonable in a number of wireless transmission scenarios (e.g. satellite conditions). We also assume that each video stream exchange lasts for $T = N\Delta$ slots, and we focus on the cases where $N \gg c$. This is also a reasonable assumption as the video exchanges often last a long period of time in many wireless transmission scenarios.

We point here that this scenario is general and is applicable to any bent-pipe topology. The model captures the scenarios in which the relay node may be a terrestrial node or a satellite node and can be adapted to different wireless scenarios by adjusting the parameters (e.g. delay in transmission depending on satellite or terrestrial scenario).

2.1.3 Cross-layer design model

The video content is transmitted using User Datagram Protocol over Real-Time Transport Protocol (RTP). The source node is equipped with a codec responsible for compression of the video content. The output rate of the codec can be reconfigured to deliver a target bit-rate. Therefore, at the source node, the output rate can be adjusted as desired. This output rate from any node $k \in \{A, B\}$ at time $t = n\Delta$ is denoted by $r_k(n)$, and it is measured in packets per time slot. The packets are further passed down to the network layer maintaining coherence with standard protocol stack. The rate of transmission of video payload is adapted to the network conditions by optimization of the QoE perceived at the end user. With the use of Real Time Control Protocol (RTCP) signaling, the source node collects the feedback information about the round-trip time from the destination. The RTCP feedback signaling provides the source node with information to re-configure the codec rate to suit the target needs.

2.1.4 Buffer and delay model

Each of the users maintains a transmission buffer of size B_c packets. The number of packets stored in transmission buffer at time $t = n\Delta$ is given by $B(n)$. The buffer is maintained using a first in, first out (FIFO) queue. Let the source node k transmit the packets at time $t = n\Delta$ at the rate $r_k(n)$. The source node is unaware of the channel link capacity $R_c(n)$. Assume that the buffer is initially empty. If $r_k(n) < R_c(n)$, all the packets are transmitted through the channel with rate $r_k(n)$. If $r_k(n) \geq R_c(n)$, then packets are transmitted with a rate of $R_c(n)$ through the channel link. From the remaining $r_k(n) - R_c(n)$ packets, $B(n) \leq B_c$ packets are stored in the transmission buffer. The remaining packets are lost. In the next time slot, the packets stored in the buffer are transmitted through the channel. These packets face a delay in reaching the destination. The same process as above is applied to the rest of the new packets

to be transmitted, but now the effective channel capacity is $R_c(n) - B(n - 1)$.

Using the above model, we will now describe a smooth adaptation method that depends on the QoE, QoS and transmission cost derived in [15] that will enable us to study the game theoretical interaction among the nodes in subsequent sections.

2.2 Generic rate adaptation

The cross layer rate adaptation model provides a framework to optimize the output codec rate of a source node with feedback from the network in the form of RTCP signals [16, 17]. We will focus on the rate adaptation based on optimizing QoE of the video obtained by the end user. The QoE can be quantified using different metrics in a spatial domain like structural symmetry (SSIM) and in temporal domain like flow continuity. For simplicity, we choose the flow continuity of the video as the measure of the QoE. The flow continuity is defined as the probability that the delay in the network falls below a threshold thereby providing a characterization of the ability of the end user to view the video without any freezing. The rate is adapted to the channel conditions as follows. The source begins the transmission at a certain initial rate. The rate is increased with time, as long as the source observes no delay. As soon as the source observes a delay, it reduces the rate to the initial value. The process is again repeated. This adaptation is shown in Fig. 3. Mathematically, any source k begins the transmission at rate $r_k(1) = \beta < R_c(1)$, where $\beta \in \mathbb{R}^+$. With each RTCP received, the rate is updated by

$$r_k(n, \alpha_k) = \begin{cases} \beta + \delta r_k(n, \alpha_k) & \text{if no delay observed} \\ \beta & \text{if delay observed} \end{cases} \quad (4)$$

where the increment in rate is

$$\delta r_k(n, \alpha_k) = [\alpha_k (n \bmod p(n)) \Delta]^2$$

such that Δ is the time interval size, $\alpha_k \in (0, \tan(\pi/2))$ is the slope of the rate adaptation curve and $p(n)$ is the time period of the waveform in Fig. 3 given by

$$p(n) = \left\lceil \frac{\sqrt{R_c(n) - \beta}}{\alpha_k \Delta} \right\rceil + 1$$

Note that this rate adaptation is done at the source node k , and it depends on the delay statistics observed at node $l \neq k$. Henceforth, we denote the player other than k as $-k$. Note that the increment in rate can be modeled using different functions like logarithmic, linear or exponential functions. We chose here a squared function having a slow start followed by a rapid growth which makes our model resemblant to TCP to some extent [16, 17].

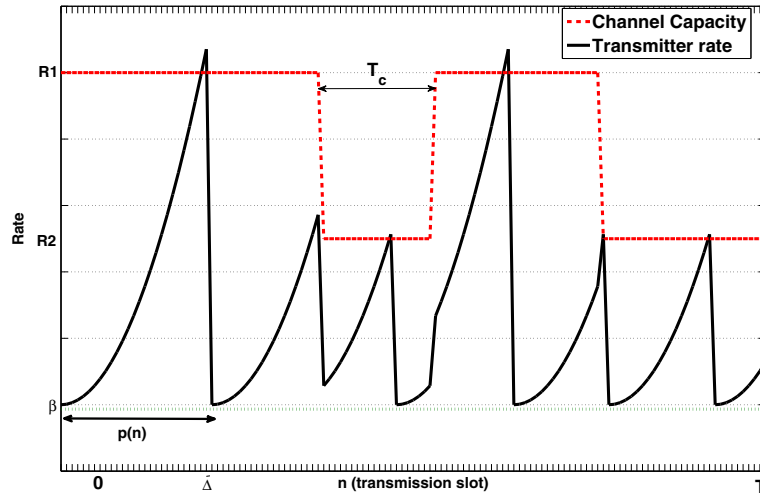


Fig. 3 Illustration of rate adaptation curve. The red and blue curves indicate the rate $r_k(n)$ and channel capacity, respectively. The angle α remains constant for all n , whereas the period $p(n)$ depends on the channel capacity

We now describe the performance metrics which will be used in this work.

2.2.1 Quality of service metric

The network utility to a node is defined as the ratio of the network capacity used to transmit packets to that node and the total network capacity. The number of packets received at node k depends on the rate of transmission at the other node $r_{-k}(n, \alpha_{-k})$. However, these packets cannot exceed the instantaneous channel capacity. Therefore, the instantaneous network utility at time $t = n\Delta$ is given by

$$\mu(R_c, r_{-k}, n, \alpha_{-k}) = \frac{\min(R_c(n), r_{-k}(n, \alpha_{-k}))}{R_c(n)}$$

where $R_c(n)$ is given by (3) and $r_{-k}(n, \alpha_{-k})$ is given by (4). We define the *Quality of Service* at node k by the averaged network utility to node k over time:

$$f_k^{\text{QoS}}(\alpha_{-k}) = \frac{1}{N} \left\{ \sum_{n=0}^N \mu(R_c(n), r_{-k}(n, \alpha_{-k})) \right\} \quad (5)$$

It can be seen that the QoS of node k depends on the rate adaptation at the other node via α_{-k} .

2.2.2 Quality of experience metric

The flow continuity is defined as the probability that the delay in the network falls below a threshold. The delay in the network leads to video packets arriving late at the end user leading to a visible freezing of the displayed video and degrading the quality of experience. For simplicity, flow continuity is chosen as the sole QoE indicator. Based on our model in Section 2.1, the delay is observed at node k

if the rate r_{-k} exceeds the channel capacity (assuming the threshold admissible delay is 0). We define a delay counter function $\varphi(\cdot)$ which determines if there is an instantaneous delay in the network or not. This function takes the instantaneous rate and the channel capacity as the input and is reset if there is no delay or gives value one otherwise:

$$\varphi(r(n, \alpha), R) = \max(0, \text{sgn}(r(n, \alpha) - R))$$

where

$$\text{sgn}(r(n, \alpha) - R) = \begin{cases} -1, & \text{if } r(n, \alpha) < R \\ 0, & \text{if } r(n, \alpha) = R \\ 1, & \text{if } r(n, \alpha) > R \end{cases}$$

Therefore, we can write the total number of delays as

$$\sum_{n=0}^N \varphi((r_{-k}(n, \alpha_{-k}), R_c(n)))$$

We can now define the *Quality of Experience* metric as the average flow continuity of the network:

$$f_k^{\text{QoE}}(\alpha_{-k}) = 1 - \left\{ \frac{\sum_{n=0}^N \varphi(R_c(n), r_{-k}(n, \alpha_{-k}))}{N} \right\} \quad (6)$$

We note that the flow continuity at k is a function of the rate adaptation gradient at the other node α_{-k} .

2.2.3 Transmission cost

In order to transmit packets, the sender node k incurs a cost of transmission due to the power usage, hardware requirements, etc. This cost is dependent on not only the

magnitude of resources used for transmission but also the gradient of increment in the rate. This is because the higher the gradient in rate increment, the more is the difference between the two consecutive instantaneous rates. This asserts a higher energy demand on the resources to make a sharper change in the rate leading to higher cost. In order to model this cost of transmission, for simplicity, we use a logarithmic function of the rate adaptation slope or α given by

$$f_k^{\text{COST}}(\alpha_k) = \log_2(1 + \arctan(\alpha_k)) \quad (7)$$

The cost of transmission at the node k is determined by his own rate adaptation gradient.

Combining QoE and QoS vs. transmission cost, we will define the player's k payoff or benefit obtained from the video exchange as the weighted sum of the three aspects

$$w_1 f_k^{\text{QoS}}(\alpha_{-k}) + w_2 f_k^{\text{QoE}}(\alpha_{-k}) - w_3 f_k^{\text{COST}}(\alpha_k)$$

where $w_i \in \mathbb{Z}$ are the respective weights and their role will be investigated via numerical simulations. Optimizing the weighted sum of objectives is a scalarization technique to solve multi-criteria optimization problems. This approach leads to good or even optimal (e.g. in convex optimization problems) trade-offs among the multiple objectives that are often opposing ones [18].

3 Game theoretical framework for QoE-driven adaptive video transmission

From the previous discussion in Section 2.1, the quality of the video obtained by a player, say A , depends on the rate of transmission of player B and vice versa. Additionally, in order to transmit the video packets to player A , player B incurs a cost and vice versa. Therefore, if the players are selfish, there is a conflict of interest as both players want to incur minimum cost (affecting the video quality of other player) and also obtain a good quality of video themselves (affected by the rate of transmission by other players). Therefore, there is an interaction arising naturally among such two nodes which is modeled using game theory.

3.1 One-shot non-cooperative game

The relevant one-shot non-cooperative game is defined by the tuple:

$$\mathcal{G}_O = \{\mathcal{P}, \{\mathcal{A}_k\}_{k \in \mathcal{P}}, \{u_k\}_{k \in \mathcal{P}}\} \quad (8)$$

in which \mathcal{P} is the set of players that selfishly maximize their own payoffs, given by u_k , by choosing the best actions in their action set \mathcal{A}_k . With the QoE-driven rate adaptation model to exchange video, we define the following component sets of the game tuple \mathcal{G}_O :

Players Set \mathcal{P} : The set of players or users is given by $\mathcal{P} = \{A, B\}$ which are the two nodes A and B that exchange the video packets.

Action Set \mathcal{A}_k : The set of actions that can be taken by the player k , where $k \in \{A, B\}$, is given by $\mathcal{A}_k = [0, \lambda \frac{\pi}{2}]$ and $\lambda \in (0, 1)$. The action chosen by the k th player from \mathcal{A}_k is denoted by α_k . Hence, the player can choose the gradient of rate adaptation as its action in order to maximize its own benefits. The factor λ ensures that the gradient is always $\alpha < \frac{\pi}{2}$ to prevent infinite increase in rate.

Payoff function u_k : The payoff function of the k th player is defined as follows:

$$u_k(\alpha_k, \alpha_{-k}) = w_1 f_k^{\text{QoS}}(\alpha_{-k}) + w_2 f_k^{\text{QoE}}(\alpha_{-k}) - w_3 f_k^{\text{COST}}(\alpha_k) \quad (9)$$

The payoff function of the k th player depends not only on his own action α_k but also on the other player's α_{-k} . This function combines jointly the video quality experienced by player k and the cost of transmitting video to the other player. The throughput and the flow continuity (given by $f_k^{\text{QoS}}(\alpha_{-k})$ and $f_k^{\text{QoE}}(\alpha_{-k})$ in (3) and (4)) affect the quality of the video experienced by player k and are both determined by the action of the other player α_{-k} . The weights $w_i \in (0, 1)$ are positive parameters that assign dimensions to the three factors such that $\sum_{i=1}^3 w_i = 1$.

A solution concept of this game is the Nash equilibrium. The Nash equilibrium (NE) of a non-cooperative game \mathcal{G} is defined as a set of actions of the players $(\alpha_k^*, \alpha_{-k}^*)$, from which no player has an incentive to deviate unilaterally. Hence, the selfish rational players can be foreseen to play the action at NE.

Proposition 1. *The unique Nash equilibrium of one-shot QoE-driven adaptive video exchange game $\mathcal{G}_O = \{\mathcal{P}, \{\mathcal{A}_k\}_{k \in \mathcal{P}}, \{u_k\}_{k \in \mathcal{P}}\}$ is given by $(\alpha_1^*, \alpha_2^*) = (0, 0)$.*

Proof. See Appendix A. \square

The NE in the one-shot game shows that the two selfish players will not exchange any data if their interaction takes place a single time. No selfish player would be willing to incur a cost to send the video when there is no guarantee of receiving anything in return. Moreover, even if the other player were to send data, rational behavior leads to the choice which minimizes the cost.

In the following, we study repeated games as a mechanism to motivate rational players to share data in the

absence of a central authority which may use other mechanisms such as pricing techniques to manipulate the Nash equilibrium [19].

3.2 Repeated games framework

We consider that the players interact repeatedly under the same conditions, i.e. the same game \mathcal{G}_O is played repeatedly. The repeated game can lead to a change in the outcome because the players can now observe the past interactions and decide accordingly upon their present action. We will first formulate the game tuple, in coherence with the definition of Section 3.1. We will consider two cases: (i) finite-horizon repeated game: the players know in advance how many times they will interact or when the game ends and (ii) infinite horizon repeated game: the players are unaware of how many times they will interact. We analyze the outcomes or game equilibria in these two cases.

Before proceeding, we remark that the assumption $N \gg c$ is required for the same game \mathcal{G}_O to be played repeatedly (see Section 2.1). The QoS and QoE terms in the players' payoff functions are empirical averages (over an N time horizon) of random quantities depending on the varying channel state. This state changes at every c temporal instances and if $N \gg c$, we can consider that these functions are approximately equal to their statistical counterparts and, thus, are good estimates as deterministic functions of α_{-k} . Otherwise, since the payoff functions would change randomly at every stage, more advanced tools such as Bayesian games would have to be used instead of repeated games.

A repeated game, in which the game \mathcal{G}_O defined in (8) is played repeatedly, is defined using the following tuple:

$$\mathcal{G}_R = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, T\}$$

whose components are defined below.

Players set \mathcal{P} : refers to the set of players $\{A, B\}$.

Strategy set \mathcal{S}_k : The strategy set of player k is different from the action set in (8) because it describes a strategic plan on how to choose an action in \mathcal{A}_k at every stage of the game and for any history of play. More precisely, let the actions taken by the players in the τ th stage be $a^{(\tau)} = (a_1^{(\tau)}, a_2^{(\tau)})$. Then the history of the game at the end of stage $t \geq 1$ is given by $h^{(t+1)} = (a^{(1)} a^{(2)} \dots a^{(t)})$. The set of all possible histories up to t is given by $\mathcal{H}^{(t)} = \{\mathcal{A}_k \times \mathcal{A}_{-k}\}^t$ where \mathcal{A}_k is the action set in (8) given by $[0, \lambda \frac{\pi}{2}]$. The strategy of a player k for T interactions is a sequence of functions

$s_k = (s_k^{(1)}, s_k^{(2)} \dots, s_k^{(T)}) \in \mathcal{S}_k$ that map each possible history to an action: $s_k^{(t)} : \mathcal{H}^{(t)} \rightarrow \mathcal{A}_k$ such that $s_k^{(t)}(h^{(t)}) = a_k^{(t)}$.

Overall payoff function v_k : The payoff function is an overall average of the payoffs obtained by player k in every stage of the game. We consider a discounted average payoff such that the player discounts the future payoffs by a factor $\delta \in (0, 1)$:

$$v_k(s) = \frac{(1 - \delta)}{(1 - \delta^T)} \sum_{t=1}^T \delta^{t-1} u_k(a^{(t)}) \quad (10)$$

where $a^{(t)}$ is the action profile at stage t induced by strategy s (i.e. $s_k^{(t)}(h^{(t)}) = a_k^{(t)} \forall k$), u_k is the one-stage payoff function defined in (9), and T is defined as the number of times the interaction takes place. It is assumed that $T > 1$.

One interesting solution in repeated games is the sub-game perfect equilibrium [10] which is a refined NE. In coherence with the NE of one-shot games, the NE of repeated games is a strategy profile from which no player gains by deviating unilaterally. However, there are some strategy profiles, which are not expected to occur because of player rationality although they are NE of the overall interaction. Hence, the sub-game perfect equilibrium region is a subset of NE. A sub-game is a repeated game starting from stage t onwards and which depends on the starting history $h^{(t)}$, and is denoted by $\mathcal{G}_R(h^{(t)})$. The final history for this sub-game is given by $h^{(T+1)} = (h^{(t)}, a^{(t)}, \dots, a^{(T)})$. The strategies and payoffs used for the sub-game are the functions of possible histories that are consistent with $h^{(t)}$. Any strategy profile s of the whole game induces a strategy $s | h^{(t)}$ on any sub-game $\mathcal{G}_R(h^{(t)})$ such that for all k , $s_k | h^{(t)}$ denotes the restriction of s_k to the histories consistent with $h^{(t)}$. A sub-game perfect equilibrium (SPE) is defined as a strategy profile $s^* = (s_1, s_2)$ such that for any stage t and any history $h^{(t)} \in \mathcal{H}^{(t)}$, the strategy $s^* | h^{(t)}$ is a NE for the relevant sub-game $\mathcal{G}_R(h^{(t)})$.

3.3 Finite horizon repeated game

We investigate the expected outcomes or SPEs for the repeated video exchange game \mathcal{G}_R assuming T is fixed and known by the two players. The only SPE of this game is given in the following proposition.

Proposition 2. *The unique sub-game perfect equilibrium of the video exchange game between two players, defined by $\mathcal{G}_R = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, T\}$, where $T < \infty$ and is known to both players is given by*

$$s_k^{(t),*} = 0, \quad \forall k \in \mathcal{P}, \forall t \in \{1, 2 \dots T\}$$

The details of the proof are provided in Appendix 2. The players are aware of the number of times the video exchange will occur; hence, the players act selfishly at each stage of the game in order to maximize their payoffs.

There is no incentive in building long-term trust in this game for any player, because it is known that the game will be played only T times. This result is similar to the repeated prisoners' dilemma [19] and is due to the fact that the players have a strictly dominant action (and action that offers strictly higher payoff than any other action irrespective of what the other player does) at every stage of the game which is $\alpha_k = 0$, for all k .

3.4 Infinite horizon repeated game

If the players' interaction is only temporary and occurs in a determined number of stages, the only rational outcome results in no video exchange at all because the player knows exactly when the interaction ends. Here, we investigate the possibility of achieving different outcomes in the case of uncertain duration or long-term interaction.

We will now study the SPE for infinite horizon repeated game. The players do not know precisely when the game will end, or equivalently, it is assumed that $T \rightarrow +\infty$. We will now identify some strategies that are SPE for such games. Note that the overall achievable SPE region for the infinite horizon repeated games is an open problem and not known in general [19].

Proposition 3. *A sub-game perfect equilibrium of the video exchange game between two players in an infinite horizon repeated game described by*

$\mathcal{G}_{\mathcal{R}} = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, +\infty\}$ *is given by*

$$s_k^{(t),*} = 0, \quad \forall k \in \mathcal{P}, \forall t \in \{1, 2, \dots, \infty\} \quad (11)$$

The proof is given in Appendix 3. This pessimistic SPE given by (11) is independent of the choice of the discount factor $\delta \in (0, 1)$. However, there are other possible SPEs. We will show that, depending upon the discount factor and other system parameters, an SPE that allows non-trivial video exchange is sustainable in the long term interaction. The intuition is similar to the infinite time horizon prisoner's dilemma [19]. In a long term, the players can build trust with one another to exchange their video (in spite of the incurred transmission cost) non-trivially and improve their received videos QoS and QoE (that depends on the opposite player's strategy) which results in overall payoff functions which are higher than the no cooperation state (0,0) for both players.

Consider the action profile (α_1^*, α_2^*) such that

$$\begin{aligned} u_1(\alpha_1^*, \alpha_2^*) &> u_1(0, 0) \\ u_2(\alpha_1^*, \alpha_2^*) &> u_2(0, 0) \end{aligned} \quad (12)$$

We focus on action profiles that provide higher payoffs than the one-shot NE for both players. Each player is willing to take the risk of paying a transmission cost in the hope that the other player will do the same and which will lead them both to higher average payoffs.

In the next proposition, we describe such SPE of the game $\mathcal{G}_{\mathcal{R}}$ which is conditional to the value of the discount factor δ .

Proposition 4. *In an infinite-horizon repeated game $\mathcal{G}_{\mathcal{R}} = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, +\infty\}$, for any agreement profile (α_1^*, α_2^*) satisfying (12), if the discount factor is bounded by*

$$1 > \delta > \delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*) \quad (13)$$

where $\delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*)$ is given by

$$\delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*) = \max_{k \in \mathcal{P}} \frac{w_3[f_k^{\text{COST}}(\alpha_k^*) - f_k^{\text{COST}}(0)]}{w_1[f_k^{\text{QoS}}(0) - f_k^{\text{QoS}}(\alpha_{-k}^*)] + w_2[f_k^{\text{QoE}}(0) - f_k^{\text{QoE}}(\alpha_{-k}^*)]} \quad (14)$$

then the following strategy is an SPE: "A player k transmits with gradient α_k^* at the first stage and continues to adapt the rate with this gradient as long as the other player adapts its rate at least by α_{-k}^* . If a defection is detected, both players stop transmitting (i.e. $\alpha = 0$) for the rest of the interaction."

We detail this proof in Appendix 4.

We remark that the inequality $0 < \delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*) < 1$ holds for any (α_1^*, α_2^*) satisfying (12) as explained in Appendix 4; therefore, (13) does not imply any additional condition on the system parameters. If an agreement point satisfies (12), then there is an admissible discount factor range within $(\delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*), 1)$ such that the agreement point (α_1^*, α_2^*) is sustainable. Intuitively, if an agreement point provides a higher utility than the one-shot NE to both players, such agreement point can be sustained.

From the above Proposition 4, we have identified the set of discount factors leading to a long-term sustainable SPE different than (0,0). The discount factor can be interpreted as the players' belief on the game to go on at every stage of the game. If the probability of the game to continue is large enough, the players develop trust and obtain better overall payoffs than minimizing their instantaneous costs.

3.5 Selection of sustainable agreement profiles

We have shown that, when the video exchange is performed repeatedly, the players can transmit the video at an agreement profile defined in (12). A natural question arises: out of all these agreement profiles, which specific profile is more likely to be selected by the players? In general, this is an open and difficult question. Also, a unified framework to tackle this problem [19] is still missing.

In this section, we present a qualitative analysis of this problem specific to our video exchange scenario. We illustrate numerically which of the achievable agreement

profiles are more likely to occur over a period of time. We consider two factors that play a role in choosing a particular agreement profiles: the tolerance index of the players and the Pareto optimality of the agreement points.

3.5.1 Tolerance index

It can be easily observed that the achievable agreement profiles do not provide equal payoffs to both the players. Some of these agreement profiles are advantageous to one of the players, and only a subset provides equal payoffs to both players.

Until now, we have assumed that every player k agrees to the action profile $(\alpha_k^*, \alpha_{-k}^*)$ without considering the payoff obtained by the other player, $u_{-k}(\alpha_k^*, \alpha_{-k}^*)$, as long as its own payoff $u_k(\alpha_k^*, \alpha_{-k}^*)$ is better than the NE. This means that the players are selfish but not malicious. However, in realistic scenarios, if the two selfish players have similar negotiation stand points, it is unlikely that they agree to a profile $(\alpha_k^*, \alpha_{-k}^*)$ that provides a huge advantage to one of the players.

In addition, there may also be cases when a player agrees to an action profile that offers a large advantage to the other player and leads to asymmetric payoffs in situations when the player wishes to obtain the video at any cost. Such a behavior is critical from the point of view of service provider providing the transmission to both the players: the provider prefers to provide just enough rate (from the other player) as desired by the player and not more, since the player is ready to settle for lower rates due to its extreme needs.

To model such behavior, we introduce the concept of tolerance index as follows:

Corollary 1. *Assuming the players (in the infinite horizon repeated game) observe both utilities at the agreement point, then an action profile (α_1^*, α_2^*) is likely to be chosen if, in addition to (12), the following condition is met*

$$|u_k^* - u_{-k}^*| < \xi \min(u_k^*, u_{-k}^*) \quad (15)$$

where $u_k^* = u_k(\alpha_k^*, \alpha_{-k}^*)$, $u_{-k}^* = u_{-k}(\alpha_k^*, \alpha_{-k}^*)$ and $\xi > 0$ represents the tolerance index.

The agreement points in (12) do not imply that the players' utilities must be equal. If the players have infinite tolerance index, then the players may agree to any action profile as long as they obtain a higher payoff than the NE. If the players have a limited tolerance index ξ , then the condition (15) restricts the previous region of agreement points as follows: the player obtaining a lower payoff will only agree to (α_1^*, α_2^*) if the difference in payoffs is lower than ξ times its own payoff. In other words, when the players exchange the video, they can tolerate a difference in their video qualities or costs which is bounded. Rational

players that have similar negotiating stand points will only agree to fair contracts.

Note that apart from this tolerance index, there exist other parameters that could be used to quantify the fairness among the different users. For example, Jain's fairness index [20] and max-min fairness index capture the fairness among all the users at the holistic level and from a centralized (system-wise) point of view. In this work, we use the tolerance index for a different reason. Indeed, our tolerance index is a user-centric one in which the individual tolerance of each user is quantized independently from the overall system.

3.5.2 Pareto optimality

Aside from the fairness criterion, Pareto optimality also plays a role in determining the profiles most likely to be chosen by rational players [19]. Pareto optimal profiles are those profiles starting from which no player can improve its own payoff without making another's payoff worse. These profiles lie on boundary of the whole achievable region. Intuitively, players will tend to agree upon the profiles which are Pareto optimal: assume that the players choose an agreement profile which is not Pareto optimal. This means that either or both players can improve their payoffs without worsening anyone's payoff. Therefore, the players will tend to agree upon the agreement profiles which are Pareto optimal such that both players can obtain the maximum possible payoff at a given discount factor and tolerance index.

4 Numerical results

We now present our simulation results to illustrate how the channel conditions modify the outcome when two players exchange video repeatedly over an undefined horizon of time. We consider the following scenario: $R_1 = 350$ kbps (a good state channel), $R_2 = 200$ kbps (a bad state channel), $T_c = 20$ s (coherence time), $T = 20000$ s (duration of each video exchange), $\beta = 150$ kbps, $\Delta = 1$, $w_1 = w_2 = 0.45$, and $w_3 = 0.1$.

4.1 Achievable agreement region

In this subsection, we present the variation in the achievable agreement points at which the players exchange videos. Given our varying channel model, the probability of a channel to be in a good state (p_G) implicitly affects these agreement points. Figure 4 illustrates the achievable agreement region of Proposition 4 as function of the varying channel conditions. We assume $\delta \rightarrow 1$ which implies a probability of the game to go on that approaches 1. When p_G tends to 1, the channel is in good conditions with high probability. This leads to a larger agreement region of payoffs exceeding the NE for both players (12). When the channel conditions worsen and p_G tends to 0, this region diminishes. This implies that the players

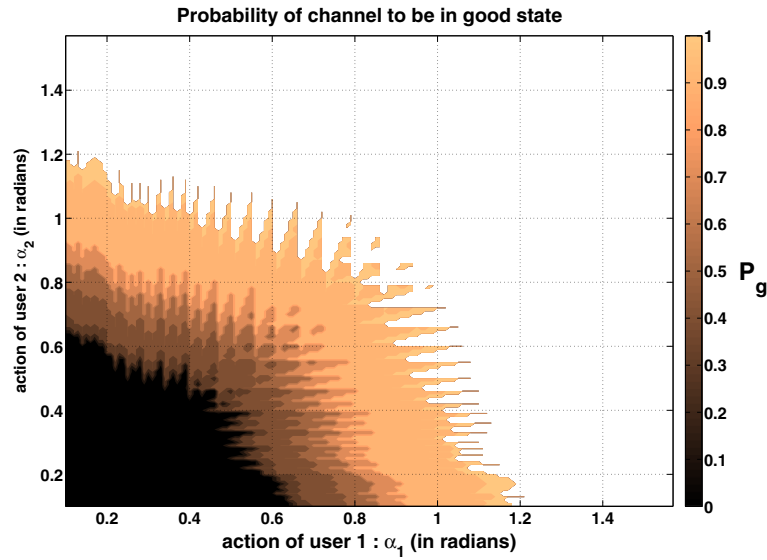


Fig. 4 Achievable region increases as channel conditions improve and there is a higher probability of channel to be in good conditions

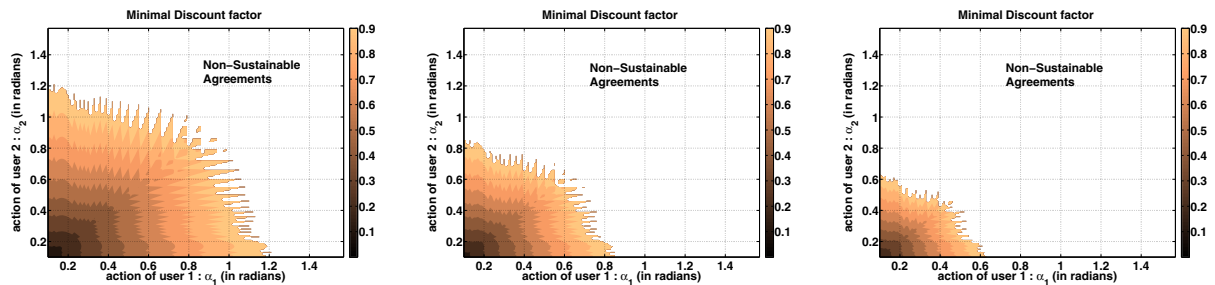
will adapt to the channel conditions faster, with a higher gradient, when the conditions are good (thereby utilizing the available resources more efficiently). However, in bad channel conditions, the players will be more cautious and will not agree to fast adaptation rates (the available resources remain under-utilized).

We remark that the boundary curve has periodical pikes that can be explained by the shape of the utility $u_k(\alpha_k, \alpha_{-k})$ as function of (α_k, α_{-k}) . For a fixed α_{-k} , this function decreases with α_k . However, this utility function is not a monotonous function of α_{-k} . This is because the utility function is composed of QoE and QoS terms aside from the transmission cost. Although, the QoE decreases with α_{-k} (due to abrupt breaks in flow continuity with a rate exceeding the channel capacity), the QoS may either increase or decrease as a function of $\{\alpha_{-k}\}$, depending on magnitude of α_{-k} . Therefore, due to a joint effect of variation of QoS and QoE with α_{-k} , the overall utility function

does not vary monotonously as a function of α_{-k} . In fact, the derivative of the utility function as a function of $\{\alpha_{-k}\}$ is a complex trigonometric function which is periodic in nature leading to the shape of the curve.

4.2 Minimal discount factor

Figure 5 illustrates the minimum discount factor necessary to achieve any agreement point depending on different channel conditions. As the achievable region decreases when passing from good to bad channel conditions (as in Fig. 5a, c), the minimum discount factor also varies proportionally. This can be observed also in Fig. 6 in which we assume that the players agree on the symmetric action profiles (or equal gradients). As the channel conditions improve (and p_G varies from 0 to 1), a lower probability of the game to continue is required. This implies that even when the probability of game to continue is low, the players have incentives in transmitting



(a) Probability of channel in good state = 1 (b) Probability of channel in good state = 0.5 (c) Probability of channel in good state = 0

Fig. 5 Variation in minimal discount factor within an achievable region for various channel conditions

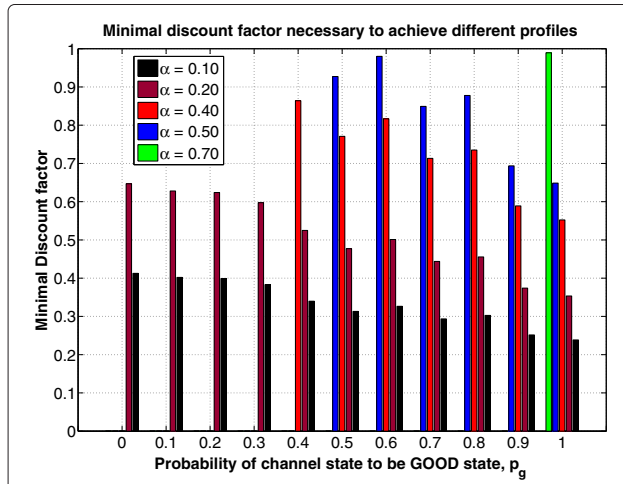


Fig. 6 Different minimal discount factor needed to achieve different symmetric action profiles for various channel conditions. The better is the channel, the lower is the discount factor needed to achieve profiles farther away from (0,0)

at a higher gradient than the NE to achieve a high utility if the channel condition is good. We observe that the players transmit at gradient higher than $\alpha = 0.7$ only in good channel conditions. In bad channel conditions, the players agree on lower gradients only. Therefore, both the channel conditions and the probability of the game to continue determine the rate of adaptation to channel conditions agreed upon by the players.

4.3 QoS vs. QoE trade-off

Assuming that the players agree only to those symmetric action profiles that are Pareto optimal, Figs. 7 and 8 show the QoS and QoE as functions of the channel conditions

and minimal discount factor. The overall QoS improves as the channel conditions improve (Fig. 7). This is due to a higher number of packets transmitted when the channel has higher average throughput capacity. The channel conditions are not very crucial in obtaining the flow continuity or QoE (Fig. 8). This is because QoE as modeled in this paper depends on the probability of rate to exceed the channel capacity leading to delays. It is independent of the magnitude of the channel capacity.

Figure 7 also shows that as the discount factor increases, the QoS increases. On the contrary, in Fig. 8, as the discount factor increases, the QoE decreases. As the discount factor increases, the probability of the game to continue also increases which allows the players to agree on higher rate adaptation gradients. This causes a higher number of packets to be transmitted per unit time which improves the overall QoS. However, higher gradients cause a higher chance of the rate to exceed the channel capacity resulting in delays and freezing of video: lower QoE. Therefore, a trade-off arises between QoS and QoE depending on the discount factor.

4.4 Tolerance index and Pareto optimality

In all previous results, we have assumed the tolerance index ξ to be infinite. In Fig. 9, we show the region of sustainable agreement points as function of the tolerance index ξ . We assume that the channel is always in a good state. When the tolerance index is low (close to 1 %), the only sustainable action profiles are symmetric. The players do not agree with unfair contracts. As the tolerance index increases, the asymmetric action profiles become sustainable and the sustainable agreement region increases. The players agree to unfair contracts, as long

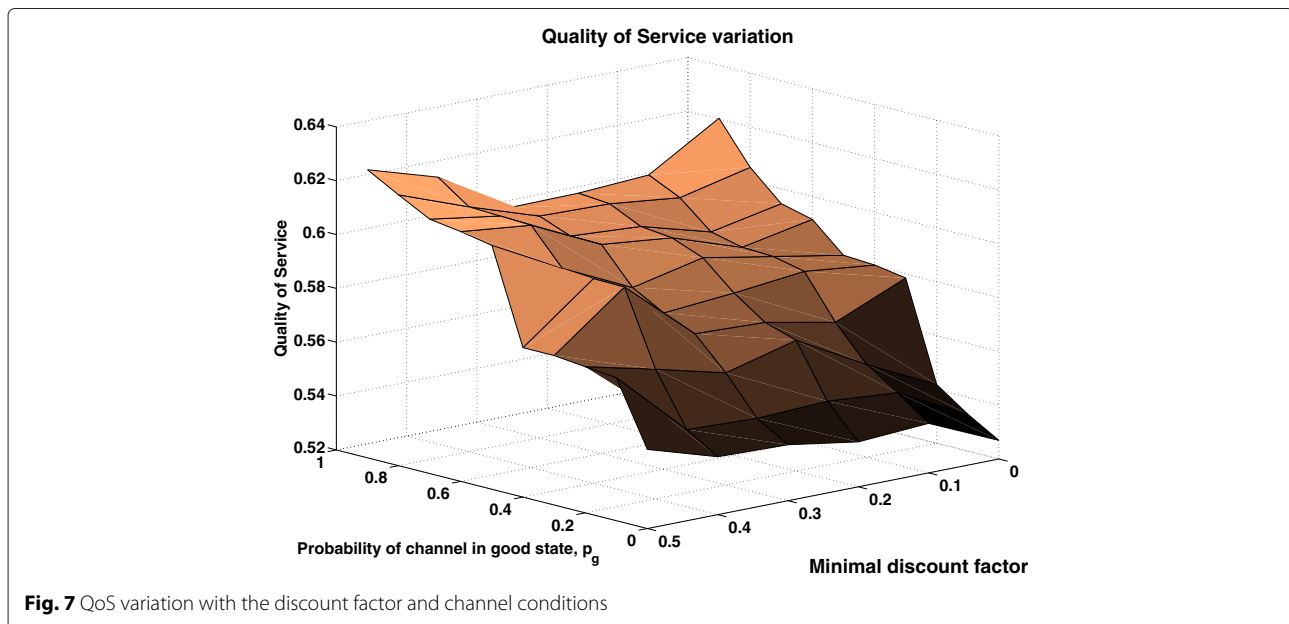


Fig. 7 QoS variation with the discount factor and channel conditions

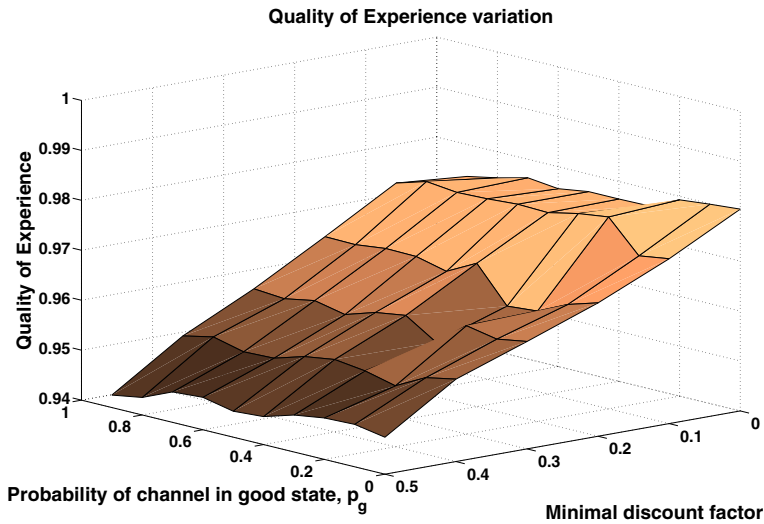


Fig. 8 QoE variation with the discount factor and channel conditions

as their own payoff is better than the NE. This behavior is favorable from service provider point of view because, even if the agreement is not fair overall, the agreement is made thereby providing just enough gradients to the players that meets their demands.

We investigate now the efficiency of any agreement point chosen by the players for a given tolerance index. Figure 10 illustrates the payoffs achieved at the sustainable agreement points. As the tolerance index increases, more asymmetric payoffs become sustainable. When $\xi = 0\%$, there is only one sustainable payoff that is Pareto optimal which plotted by the red-squared point. As the tolerance index increases, there are more sustainable payoffs that are Pareto optimal (but asymmetric).

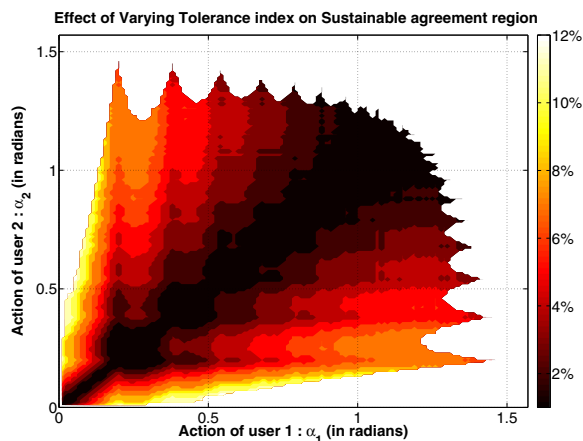


Fig. 9 Variation of sustainable agreement region for varying tolerance index. As the tolerance index increases, the non-symmetric points become achievable. With the least tolerance, the players agree to the same adaptation rates where both of them get same quality video

Another interesting observation from Fig. 10 is that when the players have a higher tolerance index, the Pareto optimal asymmetric payoffs lie on both sides of the first bisector. If the players have different leverage over each other, the player that has higher leverage (or is more powerful) can influence the payoff in its favor. For example, if player 1 has a higher leverage/power, for a tolerance index of $\xi > 0\%$, it would prefer to have a higher utility, and an agreement point below the first bisector is more likely to be chosen.

4.5 Influence of weights of QoS, QoE, and cost

We now show the dependence of the sustainable agreements' region on the different weights w_i assigned to the different factors affecting the utility in (9): QoS, QoE, and cost.

First, we focus on the scenario: $w_1 = w_2 = 0.4$ (equal weights of QoS and QoE), $p_G = 1$ (fixed channel conditions). In Figs. 11 and 4, the weight assigned to the cost is $w_3 = 0.2$ and $w_3 = 0.1$, respectively. The region of sustainable agreements is smaller in Fig. 11 than Fig. 4. This is because as the impact of cost increases, the utility u_k decreases as a function of α_k with a higher gradient. This reduces the number of points satisfying (12) and fewer agreement points offer higher payoffs than the NE. In general, the sustainable agreement region reduces when increasing the cost weight. This also implies that the players will adapt to channel capacity much slowly (at the agreement point), when the data transmission cost is high.

Second, in Fig. 12, we fix the cost weight $w_3 = 0.1$ and consider a higher weight of QoS than QoE: $w_1 = 0.6$, $w_2 = 0.3$. The region of sustainable agreements increases compared to Fig. 4. In Fig. 13, we assign a higher weight

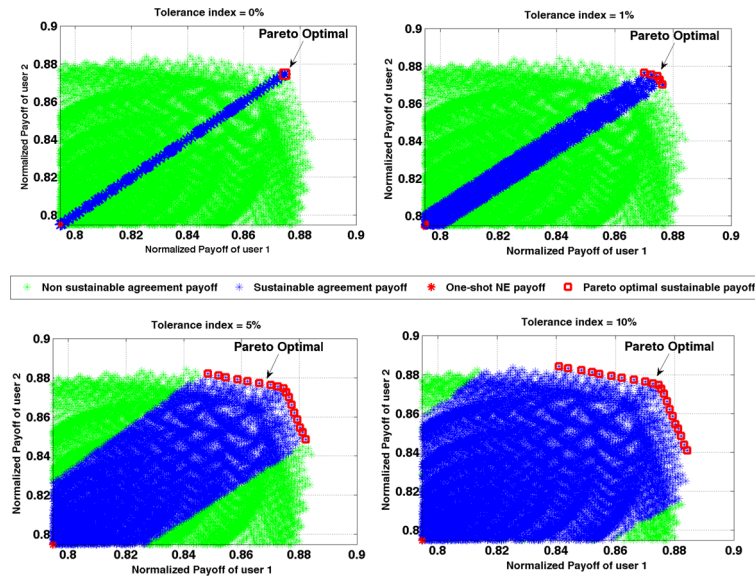


Fig. 10 Comparison of the achievable payoffs with varying tolerance index. As the tolerance index increases, more payoffs are sustainable. The number of Pareto optimal payoffs reduce as the tolerance index decreases

to QoE than QoS: $w_1 = 0.3$, $w_2 = 0.6$. We observe that the region for sustainable agreement points reduces when compared to Fig. 4. When the QoS has a higher weight, the players can agree to a faster rate adaptation which in turn provides a larger sustainable agreement region in Fig. 12. When the QoE has a higher weight, the faster rate adaptation points are not preferred by the players because they lead to a higher number of instances when the rate exceeds

channel capacity reducing the QoE. Hence, the slower rate adaptation points are sustainable in this case.

5 Conclusions

It is clear that the rapid increase in the multimedia transmission over wireless networks imposes the evolution of existing transmission protocol. As there is higher aggression among the players to compete for the existing

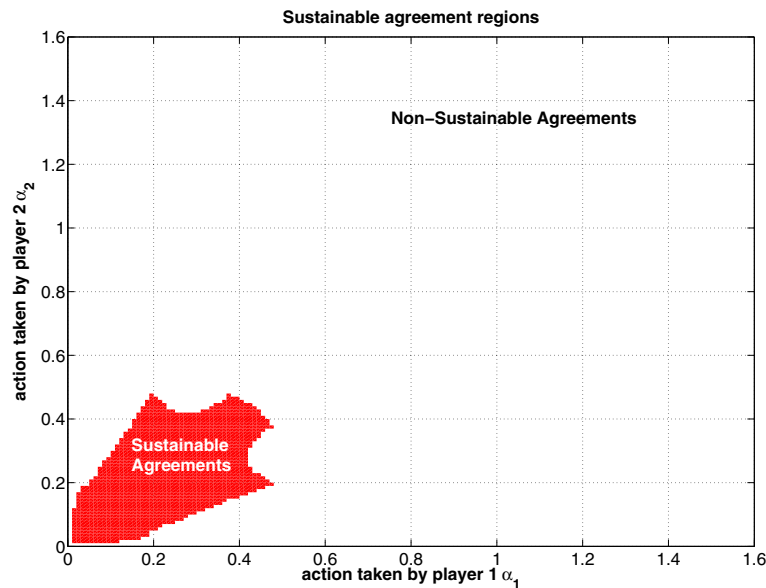


Fig. 11 The region of all agreement points (α_1^*, α_2^*) with weights assigned as $w_1 = w_2 = 0.4$ and $w_3 = 0.2$. The sustainable agreement region is reduced with a higher weightage to the cost. Due to the high cost for faster adaptation to transmit the video, the players agree to slow adaptation points

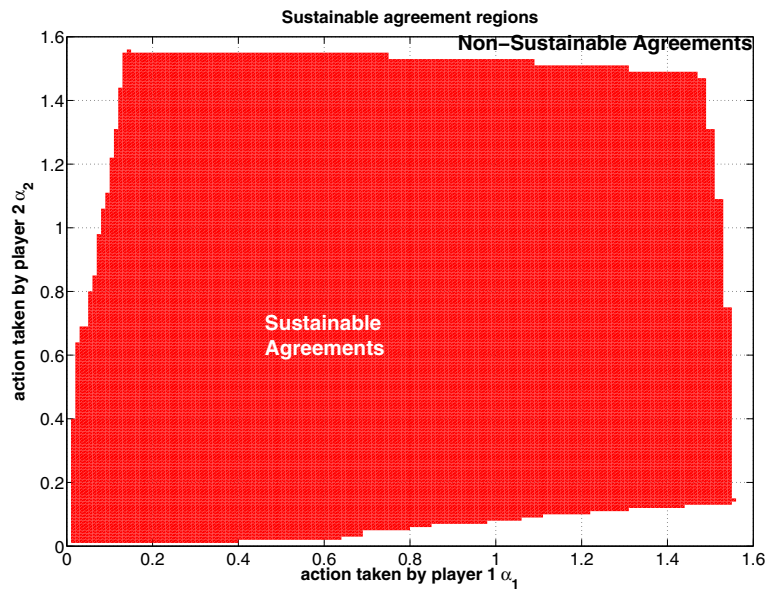


Fig. 12 Sustainable agreement points for weights $w_1 = 0.6$, $w_2 = 0.3$, and $w_3 = 0.1$. The maximum weightage is given to the QoS

resources, it is important to account for the competition within the existing protocols from the end players' point of view for stable communication systems to evolve. This paper contributes towards this aim by presenting a novel approach to model such competitions and analyzing the outcomes of these competition in a generic framework applicable to different rate adaptation protocols. It has been proved using concrete theoretical analysis and simulations that the video exchange over the wireless networks is quite sensitive to subjective considerations like

player rationality, tolerance towards other players, and the period of time for which the interaction takes place. The QoS, QoE, and cost tradeoff that we present in this work is critical towards the design of communication models for realistic setting in the presence of selfish players.

This work provides one of the primary analysis to pave the way for future research in development of robust protocols for transmission over the wireless channels. Furthermore, it is required to fine-tune this generic analysis for implementation within the scope of existent

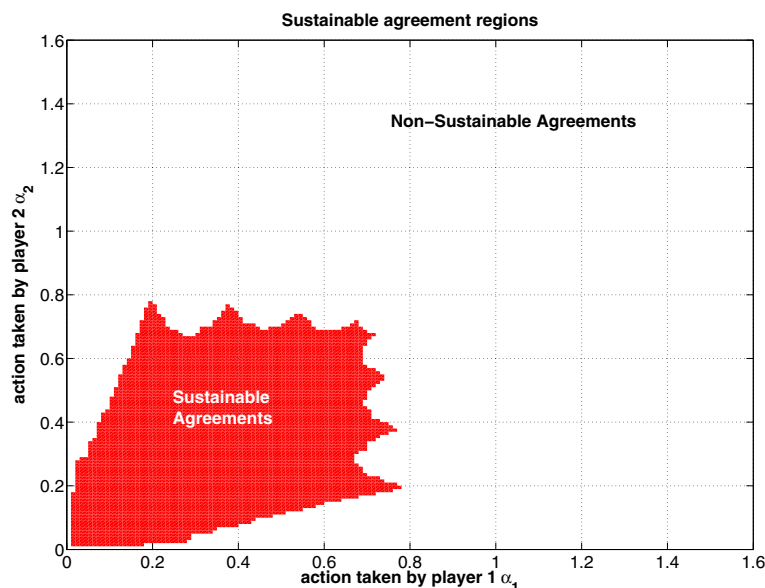


Fig. 13 Sustainable agreement points for weights $w_1 = 0.3$, $w_2 = 0.6$, and $w_3 = 0.1$. The maximum weightage is given to the QoE

protocols in order to accommodate and react to underlying trade-offs. Future lines of work also include extension of the channel models to accommodate dense networks. Another consideration for future work is the asymmetric channel conditions case, in which the throughputs of the links are different and reflective of more complex real-world scenarios. Under this perspective, future investigations could also include physical layer aspects such as modulation, and channel coding.

Appendix 1: proof of proposition 1

Consider the utility function defined in (9). Node k can maximize its utility by minimizing the cost given by (7). The cost is a logarithmic function which is an increasing function of α_k , and therefore, is minimized at $\alpha_k = 0$. The node k has no control over the action taken by the other player; therefore, the part of its utility affected by the other player's action, i.e. its QoS and QoE, cannot be controlled. In the one-shot game, when the players cannot build trust with each other, they selfishly choose the action to maximize their utility. Hence, the NE for video exchange is given by (0,0).

Appendix 2: proof of proposition 2

This proof follows by the backward induction principle [10, 19]. We consider the last stage of the game. When the game is played at the T th stage, the players are aware that they interact for the last time. They have no incentive to transmit the video for the other player while incurring a cost themselves, when there is no guarantee that the other player will transmit the video or not. Hence, their optimal strategy is $s_k^{(T),*} = 0$. Now, when the players play the game at time $T - 1$, given that in stage T they will not transmit, there is no incentive to transmit for any history $h^{(T-1)}$. Therefore, $s_k^{(T-1),*} = 0$. Following the backward induction principle, the players have no incentive to transmit at any stage in the repeated game, therefore, the SPNE is given by $s_k^{(t),*} = 0 \forall t \in \{1, \dots, T\}$. The discounted payoff is then $v_k(s^*) = 0$. Note that this result is based on the principle that a rational player will never choose an action that is strictly dominated [19]. At each stage of the game, the strategy (0, 0) is strictly dominating.

Appendix 3: proof of proposition 3

In order to prove the above proof, we cannot use backward induction as for Proposition 2, because the players are not aware of the last stage of the game. We use the one-step deviation principle to prove this SPE [10, 19]. The one-step deviation principle states that a strategy profile $s^* = (s_1^*, s_2^*)$ is an SPE if for every player k , there exists no strategy $\hat{s}_k \neq s_k^*$ such that at any stage τ and history $h^{(\tau)}$, the strategy $\hat{s}_k | h^{(\tau)}$ is a better response than $s_k^* | h^{(\tau)}$ in a sub-game $\mathcal{G}_{\mathcal{R}}(h^{(\tau)})$. This principle is based on the fact

that if there is a strategy which offers the incentive to a player to deviate at a single stage in the game then the initial strategy is not a SPE. However, if the player has no incentive for any deviation from its current strategy at any stage, such a strategy is an SPE.

To use the one-step deviation principle, firstly, we prove that the stage payoffs in (9) are uniformly bounded. Consider the expression in (9): $u_k(\alpha_k, \alpha_{-k}) = w_1 f_k^{\text{QoS}}(\alpha_{-k}) + w_2 f_k^{\text{QoE}}(\alpha_{-k}) - w_3 f_k^{\text{COST}}(\alpha_k)$. The first term can be easily written as (using (3))

$$\frac{\beta}{2} \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \leq f_k^{\text{QoS}}(\alpha_{-k}) \leq 1$$

This is due to the fact that $\beta \leq \min(R, r_k(n))$ because $\beta < R_1, R_2$ by assumption. Similarly, the second term can be written as (using (4))

$$0 \leq f_k^{\text{QoE}}(\alpha_{-k}) \leq 1$$

This follows the fact that $0 \leq \max(0, \text{sgn}(x)) \leq 1$. The third term can be written as (using (7))

$$0 < f_k^{\text{COST}}(\alpha_k) < \log \left(1 + \frac{\pi}{2} \right)$$

This is due to the limit on $\alpha_k \in [0, \tan(\frac{\pi}{2})]$. Using the above limits, we can write loose bounds of stage payoffs as

$$w_1 \frac{\beta}{2} \left(\frac{1}{R_1} + \frac{1}{R_2} \right) - w_3 \log \left(1 + \frac{\pi}{2} \right) < u_k(\alpha_k, \alpha_{-k}) \leq w_1 + w_2$$

Since $w_i \in [0, 1]$, hence, the stage payoffs are uniformly bounded.

Secondly, we will evaluate if any deviation from the profile in (11) offers a profit to one of the players. Let us assume that the player k deviates from the strategy s_k^* at stage τ and history $h^{(\tau)}$ with the strategy $\hat{s}_k^{(\tau)}(h^{(\tau)}) = \alpha_k \in (0, \lambda \frac{\pi}{2}]$ and from then on conforms again to the strategy s_k^* such that $\hat{s}_k^{(t)} = s_k^{(t),*}$ for all $t > \tau$. This implies that the player k transmits video with the rate adaptation gradient $\alpha_k > 0$, at an intermediate game played at the τ th stage and then conforms to $\alpha_k = 0$ for the rest of the game. The player k therefore incurs some extra cost of transmission at the τ th stage given by (7) which leads to $u_k(\alpha_k, 0) < u_k(0, 0)$. Therefore, the payoff $v_k(\hat{s}_k | h^{(\tau)}, s_{-k}^* | h^{(\tau)}) < v_k(s_k^* | h^{(\tau)}, s_{-k}^* | h^{(\tau)})$. Hence, the player k has no incentive in sending any video packets at any stage of the game, thereby making (11) an SPE.

Appendix 4: proof of proposition 4

We will once again use the one-step deviation principle to prove the SPE. Consider the players following the agreement profile (α_1^*, α_2^*) satisfying (12). Let us consider two cases: (i) case 1: there has been no deviation from (α_1^*, α_2^*) and (ii) case 2: there has been a deviation from (α_1^*, α_2^*) and now both players are using the threat point (0, 0).

Case 1. Let no player deviate from the agreement till the stage τ . Let, at stage $\tau + 1$, the player k deviate from the strategy and transmits at $\tilde{s}_k^{(\tau+1)} = \bar{\alpha}_k \neq \alpha_k^*$. There are two sub-cases. (1) First, we consider the case when $\bar{\alpha}_k < \alpha_k^*$. In this case, from stage $\tau + 2$ onwards, the node A conforms to the initial strategy again. Since there has been a deviation, conforming to the strategy implies $\tilde{s}_k^{(t)} = 0$ for all $t \geq \tau + 2$. We calculate the discounted payoffs in two cases. Firstly, in case there is no deviation, the discounted payoff is given by

$$v_k(s^*) = \frac{(1-\delta)}{(1-\delta^T)} \sum_{t=1}^T \delta^{t-1} u_k(\alpha_k^*, \alpha_{-k}^*) = u_k(\alpha_k^*, \alpha_{-k}^*) \quad (16)$$

Secondly, in case there is a deviation by player k at stage $\tau + 1$, the discounted payoff is given by $v_k(\tilde{s}) =$

$$\frac{(1-\delta)}{(1-\delta^T)} \left(\sum_{t=1}^{\tau} \delta^{t-1} u_k(\alpha_k^*, \alpha_{-k}^*) + \delta^{\tau} u_k(\bar{\alpha}_k, \alpha_{-k}^*) + \sum_{t=\tau+2}^T \delta^{t-1} u_k(0, 0) \right)$$

where $\bar{\alpha}_k < \alpha_k^*$. We simplify the above expression as

$$\begin{aligned} v_k(\tilde{s}) &= \frac{(1-\delta)}{(1-\delta^T)} \left(u_k(\alpha_k^*, \alpha_{-k}^*) \left[\frac{1-\delta^{\tau}}{1-\delta} \right] \right. \\ &\quad \left. + \delta^{\tau} u_k(\bar{\alpha}_k, \alpha_{-k}^*) + u_k(0, 0) \left[\sum_{t=1}^T \delta^{t-1} - \sum_{t=1}^{\tau+1} \delta^{t-1} \right] \right) \\ &\times v_k(\tilde{s}) = \frac{(1-\delta)}{(1-\delta^T)} \left(u_k(\alpha_k^*, \alpha_{-k}^*) \left[\frac{1-\delta^{\tau}}{1-\delta} \right] \right. \\ &\quad \left. + \delta^{\tau} u_k(\bar{\alpha}_k, \alpha_{-k}^*) + \left(u_k(0, 0) \left[\frac{1-\delta^T}{1-\delta} - \frac{1-\delta^{\tau+1}}{1-\delta} \right] \right) \right) \end{aligned}$$

Applying the limit $T \rightarrow \infty$, $\delta^T \rightarrow 0$. Further, taking the fractional term into the brackets, we get $v_k(\tilde{s}) =$

$$\begin{aligned} &(u_k(\alpha_k^*, \alpha_{-k}^*)(1-\delta^{\tau}) + (1-\delta)\delta^{\tau} u_k(\bar{\alpha}_k, \alpha_{-k}^*) \\ &\quad + u_k(0, 0)\delta^{\tau+1}) \end{aligned} \quad (17)$$

The discounted payoff of the strategy with one-step deviation should be lesser than the strategy with no deviation, for the latter to be SPE. Therefore, we identify the δ such that

$$v_k(s^*) > v_k(\tilde{s})$$

Inserting the values from (16) and (17), we get

$$\begin{aligned} u_k(\alpha_k^*, \alpha_{-k}^*) &> (u_k(\alpha_k^*, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)\delta^{\tau} + \delta^{\tau} u_k(\bar{\alpha}_k, \alpha_{-k}^*) \\ &\quad - \delta^{\tau+1} u_k(\bar{\alpha}_k, \alpha_{-k}^*) + \delta^{\tau+1} u_k(0, 0)) \end{aligned}$$

By (12), $u_k(\alpha_k^*, \alpha_{-k}^*) > 0$ because $u_k(0, 0) > 0$ for sustainable points¹, we get

$$0 > \delta^{\tau} (-u_k(\alpha_k^*, \alpha_{-k}^*) + u_k(\bar{\alpha}_k, \alpha_{-k}^*) - \delta u_k(\bar{\alpha}_k, \alpha_{-k}^*) + \delta u_k(0, 0))$$

Since $\delta > 0$, therefore,

$$(u_k(\alpha_k^*, \alpha_{-k}^*) - u_k(\bar{\alpha}_k, \alpha_{-k}^*) + \delta u_k(\bar{\alpha}_k, \alpha_{-k}^*) - \delta u_k(0, 0)) > 0$$

Consequently,

$$[u_k(\alpha_k^*, \alpha_{-k}^*) - u_k(\bar{\alpha}_k, \alpha_{-k}^*)] + \delta [u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)] > 0$$

Collecting the terms and rewriting, we get

$$\delta > \frac{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)]}$$

Therefore, the player k has no incentive to deviate to any $\bar{\alpha}_k < \alpha_k^*$ with the above discount factor condition in order to decrease its cost. Under the following sufficient condition on the discount factor

$$1 > \delta > \max_{k \in \mathcal{P}, \bar{\alpha}_k < \alpha_k^*, \bar{\alpha}_k \in \mathcal{A}_k} \frac{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)]} \quad (18)$$

we can see that the discounted payoff $v_k(\tilde{s})$ is less than $v_k(s^*)$. Additionally, it can be seen from (12), that $[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)] < [u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)]$. Therefore, $[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)] / [u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)] < 1$. So, condition (18) does not imply any supplementary condition on system parameters. To further simplify the expression, we find the value of $\bar{\alpha}_k$ which maximizes the expression in (18). Using (9) and (7), we know that $\frac{du_k(\bar{\alpha}_k, \alpha_{-k}^*)}{d\bar{\alpha}_k} < 0$. In other words, the utility of player k monotonically decreases with increasing $\bar{\alpha}_k$ for a fixed α_{-k} . We will prove that the condition in (18) is strictly decreasing with increasing $\bar{\alpha}_k$, and reaches a maximum at $\bar{\alpha}_k = 0$. Taking the derivative of right hand side in (18), we get

$$\begin{aligned} \frac{d}{d\bar{\alpha}_k} \left\{ \frac{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)]} \right\} &= \left(\frac{du_k(\bar{\alpha}_k, \alpha_{-k}^*)}{d\bar{\alpha}_k} \right) \\ &\times \frac{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)] - [u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)]^2} \end{aligned}$$

With $\frac{du_k(\bar{\alpha}_k, \alpha_{-k}^*)}{d\bar{\alpha}_k} < 0$, and the condition (12), the above expression is strictly negative. Hence, the lower limit of δ in (18) is strictly decreasing with increasing $\bar{\alpha}_k$. Therefore, it maximizes at minimum value of $\bar{\alpha}_k$ which is 0. Thus,

$$\begin{aligned} \max_{k \in \mathcal{P}, \bar{\alpha}_k < \alpha_k^*, \bar{\alpha}_k \in \mathcal{A}_k} \left\{ \frac{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(\bar{\alpha}_k, \alpha_{-k}^*) - u_k(0, 0)]} \right\} &= \\ \max_{k \in \mathcal{P}} \left\{ \frac{[u_k(0, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(0, \alpha_{-k}^*) - u_k(0, 0)]} \right\} \end{aligned}$$

Inserting the values of u_k from (9) we get,

$$\delta > \frac{-[w_3 f_k^{\text{COST}}(0) - w_3 f_k^{\text{COST}}(\alpha_k^*)]}{[w_1 f_k^{\text{QoS}}(\alpha_{-k}^*) + w_2 f_k^{\text{QoE}}(\alpha_{-k}^*) - w_1 f_k^{\text{QoS}}(0) - w_2 f_k^{\text{QoE}}(0)]}$$

We have the limits of δ as

$$1 > \delta > \delta_{\text{asym}}^{\min}(\alpha_k^*, \alpha_{-k}^*)$$

where $\delta_{\text{asym}}^{\min}(\alpha_k^*, \alpha_{-k}^*)$ is given by

$$\max_{k \in \mathcal{P}} \frac{w_3 [f_k^{\text{COST}}(\alpha^*) - f_k^{\text{COST}}(0)]}{w_1 [f_k^{\text{QoS}}(0) - f_k^{\text{QoS}}(\alpha_{-k}^*)] + w_2 [f_k^{\text{QoE}}(0) - f_k^{\text{QoE}}(\alpha_{-k}^*)]}$$

Now we consider the the second sub-case.

(2) Let us assume that $\bar{\alpha}_k > \alpha_k^*$. This is a trivial case because transmitting at one stage at $\bar{\alpha}_k$ is not a deviation from the strategy as the player still transmits at least α_k^* . However, for completeness, we provide the analysis. Now, the player k conforms to the agreement point until the stage τ . At $\tau + 1$, it deviates from agreement point and transmits at an $\bar{\alpha}_k$ greater than α_k^* . From $\tau + 2$, it again conforms to the strategy. Conforming to the strategy implies sending at α_k^* . In this case, the discounted payoff without deviation, given by $u_k(\alpha_k^*, \alpha_{-k}^*)$, is strictly greater than the discounted payoff with deviation. This is because, at stage $\tau + 1$, the payoff $u_k(\bar{\alpha}_k, \alpha_{-k}^*) < u_k(\alpha_k^*, \alpha_{-k}^*)$ for any $\bar{\alpha}_k > \alpha_k^*$. Further, from stage $\tau + 2$ onwards, the payoff at each stage after deviation is equal to the payoff at each stage without deviation, i.e. $u_k(\alpha_k^*, \alpha_{-k}^*)$. Therefore, overall, the discounted payoff without deviation is strictly greater than discounted payoff with deviation. We now consider case 2.

Case 2. We now consider the case when there has been a deviation from the agreement point $(\alpha_k^*, \alpha_{-k}^*)$ and now both players are using the threat point $(0, 0)$. We would now examine if it is possible for any player to deviate from this strategy and play $\alpha_k > 0$. We note that $u_k(\alpha_k, 0) < u_k(0, 0)$. Hence, if any player deviates at any intermediate stage from $(0, 0)$ and then conforms to the strategy by playing $(0, 0)$, the stage payoff at the deviation will be lesser than the stage payoff if there was no deviation. Therefore, the overall payoff without deviation would be higher than the payoff with deviation. Therefore, the player has no incentive to deviate from $(0, 0)$.

In addition, to prove that $0 < \delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*) < 1$, we write $\delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*)$ as

$$\max_{k \in \mathcal{P}} \left\{ \frac{[u_k(0, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(0, \alpha_{-k}^*) - u_k(0, 0)]} \right\}$$

From (12), we have $u_k(\alpha_k^*, \alpha_{-k}^*) > u_k(0, 0)$. It follows that $-u_k(\alpha_k^*, \alpha_{-k}^*) < -u_k(0, 0)$. Adding $u_k(0, \alpha_{-k}^*)$ on both sides we get $u_k(0, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*) < u_k(0, \alpha_{-k}^*) -$

$u_k(0, 0)$. Thus, the inequality $\frac{[u_k(0, \alpha_{-k}^*) - u_k(\alpha_k^*, \alpha_{-k}^*)]}{[u_k(0, \alpha_{-k}^*) - u_k(0, 0)]} < 1$ holds. Both, the numerator and denominator, are positive because $u_k(0, \alpha_{-k}^*) > u_k(\alpha_k^*, \alpha_{-k}^*)$ and $u_k(0, \alpha_{-k}^*) > u_k(0, 0)$. Consequently, we obtain $0 < \delta_{\text{asym}}^{\min}(\alpha_1^*, \alpha_2^*) < 1$. This completes the proof.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Telecommunications and Systems Engineering, Autonomous University of Barcelona, Cerdanyola del Vallès, Spain. ²ETIS / ENSEA - UCP - CNRS, Cergy-Pontoise, France. ³Inria, Rocquencourt, France. ⁴CA Strategic Research, CA Technologies, Barcelona, Spain.

Received: 16 August 2015 Accepted: 24 February 2016

Published online: 05 March 2016

References

1. RKP Mok, EWW Chan, RKC Chang, in *Integrated Network Management*, ed. by N Agoulmine, C Bartolini, T Pfeifer, and D O'Sullivan. Measuring the quality of experience of HTTP video streaming (IEEE, 2011), pp. 485–492
2. AB MacKenzie, LA DaSilva, *Synthesis Lectures on Communications Morgan & Claypool Publishers*, (2006)
3. S Lasaulce, H Tembine, *Game Theory and Learning for Wireless Networks: Fundamentals and Applications*, 1st ed. (Academic Press, 2011)
4. EV Belmega, S Lasaulce, M Debbah, Power allocation games for mimo multiple access channels with coordination. *IEEE Trans. Wirel. Commun.* **8**(6), 3182–3192 (2009)
5. YE Sagduyu, A Ephremides, in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on. Vol. 4*. Power control and rate adaptation as stochastic games for random access (IEEE, 2003), pp. 4202–4207
6. J Ye, M Hamdi, in *Global Telecommunications Conference (GLOBECOM 2011)*. Priority-Based Rate Adaptation Using Game Theory in Vehicular Networks (IEEE, Houston, TX, USA, 2011), pp. 1–6. doi:10.1109/GLOCOM.2011.6133998
7. TA Trinh, S Molnár, in *Quality of Service in the Emerging Networking Panorama*. A game-theoretic analysis of tcp vegas (Springer, Berlin Heidelberg, 2004), pp. 338–347
8. H Ren, MH Meng, Game-theoretic modeling of joint topology control and power scheduling for wireless heterogeneous sensor networks. *IEEE Trans. Autom. Sci. Eng.* **6**(4), 610–625 (2009)
9. MA Vázquez-Castro, F Perez-Fontan, LMS Markov model its use for power control error impact analysis on system capacity. *IEEE J. Sel. Areas Commun.* **20**(6), 1258–1265 (2002)
10. EV Belmega, L Sankar, HV Poor, Enabling data exchange in two-agent interactive systems under privacy constraints. *IEEE J. Sel. Top. Signal Process.* **9**(7), 1285–1297 (2015)
11. M Le Treust, S Lasaulce, A repeated game formulation of energy-efficient decentralized power control. *IEEE Trans. Wirel. Commun.* **9**(9), 2860–2869 (2010)
12. A Akella, S Seshan, R Karp, S Shenker, C Papadimitriou, Selfish behavior and stability of the internet: a game-theoretic analysis of TCP. *SIGCOMM Comput. Commun. Rev.* **32**(4), 117–130 (2002)
13. S Gupta, EV Belmega, MA Vázquez-Castro, in *Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC), 7th*. Game theoretical analysis of the tradeoff between QoE and QoS over satellite channels, (Livorno, 2014), pp. 24–31
14. E Lutz, A Markov model for correlated land mobile satellite channels. *Int. J. Satell. Commun.* **14**, 333–339 (1996). Wiley Subscription Services, Inc., A Wiley Company
15. S Gupta, MA Pimentel-Niño, MA Vázquez-Castro, in *Mosharaka 3rd International Conference on Wireless Communications and Mobile Computing (MIC-WCMC 2013)*. Joint network coded-cross layer optimized video streaming over relay satellite channel, (Valencia, 2013), pp. 14–16
16. MA Pimentel-Niño, MA Vázquez-Castro, H Skinnemoen, in *30th AIAA International Communications Satellite Systems Conference*. Optimized

ASMIRA advanced QoE video streaming for mobile satellite communications systems, (Ottawa, 2012)

17. A Afanasyev, N Tilley, P Reiher, L Kleinrock, Host-to-host congestion control for TCP. *IEEE Commun. Surv. Tutor.* **12**(3), 304–342 (2010). Third Quarter. doi:10.1109/SURV.2010.042710.00114
18. S Boyd, L Vandenberghe, *Convex optimization*. (Cambridge University Press, New York, NY, USA, 2009)
19. D Fudenberg, J Tirole, *Game Theory*. (MIT Press, Cambridge, MA, 1991)
20. R Jain, DM Chiu, W Hawe, *A quantitative measure of fairness and discrimination for resource allocation in shared computer system*, vol. 38. (Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA, 1984)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com